

# Dynamics of geodesic flows with random forcing on Lie groups with left-invariant metrics

W. Hu and V. Šverák

University of Minnesota

## Abstract

We consider stochastic perturbations of geodesic flow for left-invariant metrics on finite-dimensional Lie groups and study the Hörmander condition and some properties of the solutions of the corresponding Fokker-Planck equations.

## 1 Introduction

Our motivation for this paper comes from the problem of turbulent mixing. However, instead of studying the motion of fluids, which can be mathematically described by trajectories in the group of diffeomorphisms of the domain containing the fluid (as pointed out by V. I. Arnold [1]), we will study its finite-dimensional version when the diffeomorphism group is replaced by a finite-dimensional Lie group  $G$ . We equip  $G$  with a left-invariant metric, and consider stochastically perturbed geodesic flows. This situation may admittedly be somewhat removed from important phenomena in the real flows related to very high (or even infinite) dimensionality of the phase-spaces relevant there (at least if do not take the dimension of  $G$  as large parameter), but it does retain an important feature: the amplification of the stochastic effects by the non-linearity. A well-known example of this effect in the context of 2d incompressible fluids has been established in a seminal paper by Hairer-Mattingly [8], where ergodicity for stochastically forced 2d Navier-Stokes equation was proved for degenerate forcing, under optimal assumptions. There are at least two important themes involved in this result. One might perhaps be called algebraic, and involves calculations of Lie algebra hulls related to the Hörmander hypoellipticity condition [12, 6]. The other one belongs to Analysis/Probability and deals with consequences of the Hörmander condition (which are of course already of great interest in finite dimension) in the infinite-dimensional setting (under suitable assumptions). In the finite-dimensional models we consider in this paper the analysis component is much simpler, although there still are many non-trivial and interesting issues related to various aspects of hypoelliptic operators, such as the domains of positivity of the of fundamental solutions and convergence to equilibria.

Our focus here will be on the algebraic part. Roughly speaking, we will be interested in algebraic conditions which imply the Hörmander condition, ergodicity and convergence to equilibria. The stochastic forces will be essential for this, but it is interesting to try in some sense to minimize the “amount” of stochasticity which is needed.

One can of course also study the ergodicity of the non-stochastic dynamics, but we have nothing new to say about this notoriously difficult problem.

We consider two different types of models. The first one might be called the Langevin-type

perturbation of the geodesic flow. It is related to the stochastic equation

$$\ddot{x} + \nu \dot{x} = \xi, \quad (1.1)$$

where  $\xi$  is a random force, which is “degenerate”, in the sense that it acts only in a few directions. On a group  $G$  with a left-invariant metric (and under suitable assumptions on  $\xi$ ) one can employ symplectic reduction and obtain an equation

$$\dot{z}^k = q^k(z, z) - \nu z^k + \sigma_l^k \dot{w}^l, \quad (1.2)$$

in the Lie algebra  $\mathfrak{g}$  of the group, where we sum over repeated indices,  $k$  runs from 1 to the dimension of the group,  $l$  runs from 1 to the dimension of the noise (which can be 1),  $w^l$  are independent standard Wiener processes, and the equation

$$\dot{z} = q(z, z) \quad (1.3)$$

is the Euler-Arnold equation in  $\mathfrak{g}$  as established in [1]. For this model we determine an algebraic condition on  $q$  which is necessary and sufficient for the Hörmander condition for the corresponding Fokker-Planck equation to be satisfied in the cotangent space  $T^*G$ , see Theorem 3.1. For a compact group  $G$  this condition implies ergodicity, and the projection of the ergodic measure to  $G$  is the Haar measure. This means that the (stochastically perturbed) geodesic flow will visit all points on the group with the same probability (with respect to the Haar measure). We note that in the setting of the left-invariant metrics on a group this can never be the case without forcing, due to known conserved quantities one gets from Noether’s theorem.

For our next group of models we take a compact manifold  $Z \subset \mathfrak{g}$  which is invariant under the flow of (1.3) and consider

$$\dot{z} = q(z, z) + \xi, \quad (1.4)$$

where  $\xi$  schematically stands for random forcing induced by the Brownian motion in  $Z$  with respect to a natural Riemannian metric. One example we have in mind - in the co-tangent bundle  $T^*G$  picture<sup>1</sup> - is the intersection of a co-adjoint orbit and an energy level. The manifold  $Z$  can have much lower dimension than  $G$ . This situation may in fact be a fairly realistic description of a motion with random perturbations in which the quantities defining  $Z$  are monitored and kept close to constant values by some control mechanism. When combined with random perturbations, such control might easily induce random drift along the surface defined by specified values of the controlled quantities. (A more concrete mathematical process is described in subsection 3.3.) Together with the equation

$$a^{-1}\dot{a} = z, \quad (1.5)$$

the stochastic equation (1.4) gives a stochastic equation in  $G \times Z$ . In this situation we again determine an algebraic condition on  $Z$  which is necessary and sufficient for the Fokker-Planck equation in  $G \times Z$  associated to (1.4) to satisfy the Hörmander condition, see Theorem 3.2, and, when the condition is satisfied, establish ergodicity and convergence to equilibrium. For compact  $G$  the ergodic measure is given by a product of the Haar measure on  $G$  and an invariant measure on  $Z$ .

In the case of a non-compact  $G$  and a compactly supported initial condition for the Fokker-Planck equation the behavior will of course be different, and we illustrate what one might expect by an explicit calculation for  $G = \mathbf{R}^n$  and a one-dimensional manifold  $Z$ , see Proposition 3.1.

---

<sup>1</sup>We can of course go back and forth between  $TG$  and  $T^*G$  with the help of the metric.

The themes above have strong connections to control theory. In addition to the remark above about interpreting  $Z$  as a “control surface”, there is another connection via the Stroock-Varadhan Theorem [16]. Roughly speaking, instead of random forcing  $\xi$  one can consider forcing by control and ask which states can be reached (and how efficiently). For an introduction to control theory see for example [5].

## 2 Preliminaries

### 2.1 Basic notation and setup

Let  $G$  be a Lie group. Its elements will be denoted by  $a, b, \dots$ . We will denote by  $\mathfrak{g}$  and  $\mathfrak{g}^*$  respectively its Lie algebra and its dual. Let  $e_1, \dots, e_n$  be a basis of  $\mathfrak{g}$  and let  $e^1, \dots, e^n$  be its dual basis in  $\mathfrak{g}^*$ , determined by  $\langle e^i, e_j \rangle = \delta_j^i$ . We assume that a metric tensor with coordinates  $g_{ij}$  in our basis is given on  $\mathfrak{g}$ . In what follows we will mostly use the standard formalism of orthonormal frames and assume that  $g_{ij} = \delta_{ij}$ , which can of course always be achieved by a suitable choice of the original basis, although sometimes it may be useful not to normalize  $g_{ij}$  this way, so that other objects could be normalized instead. When  $g_{ij} = \delta_{ij}$ , we can then identify vectors with co-vectors without too much notation and write  $|x|^2$  for the square of the norm of an  $x \in \mathfrak{g}$  or  $x \in \mathfrak{g}^*$  given by the metric tensor. However, we will try to avoid relying on this normalization too much and many of our formulae will be independent of it. In such situations we will use the classical convention of using upper indices for vectors and lower indices for co-vectors, with the usual conventions  $y_k = g_{kl}y^l$  and  $y^k = g^{kl}y_l$ , where  $g^{kl}$  is the inverse matrix of  $g_{kl}$ . In this notation we can for example write  $|y|^2 = y_k y^k$ .

The various objects on  $\mathfrak{g}$  and  $\mathfrak{g}^*$  can be transported to  $T_a G$  and  $T_a^* G$  for any  $a \in G$  in the standard way, by using the left translation  $b \rightarrow ab$ . The resulting frame of vectors fields on  $G$  (or 1-forms), will still be denoted by  $e_1, \dots, e_n$ .

We can then consider  $G$  as a Riemannian manifold. The left translations  $b \rightarrow ab$  are more or less by definition isometries of the manifold. They obviously act transitively on  $G$ , and hence  $G$  is a homogeneous Riemannian manifold.

The relevance of this construction for the mechanics of fluids and rigid bodies was pointed out in Arnold’s paper [1] already mentioned above. The main point is that for fluids and rigid bodies the configuration space of the corresponding physical system is naturally given by a group (which however is infinite-dimensional for fluids), and the kinetic energy given a natural metric tensor on it. We refer the reader to the book by Arnold and Khesin [2] for a deeper exposition of these topics and additional developments.

### 2.2 The symplectic structure in $T^*G$ in left-invariant frames

The cotangent space  $T^*G$  is the canonical phase space for describing the geodesic flow in  $G$  via the Hamiltonian formalism. For a group  $G$  with a left-invariant metric the space  $T^*G$  can be identified with  $G \times \mathfrak{g}^*$  by using the frame  $e_1, \dots, e_n$  on  $G$ :

$$(a, y) \in G \times \mathfrak{g}^* \quad \rightarrow \quad y_k e^k(a) \in T_a^* G, \quad (2.1)$$

where  $e^1, \dots, e^n$  is the frame in  $T^*G$  which is dual to  $e_1, \dots, e_n$ . Here and in what follows we use the standard convention of summing over repeated indices. The “coordinates” in  $T^*G$  given by  $(a, y)$  are convenient for calculations, and will be freely used in what follows. Note that the prolongation of the action  $a \rightarrow ba$  of  $G$  on itself to  $T^*G$  has a very simple form in the  $(a, y)$  coordinates:

$$(a, y) \rightarrow (ba, y), \quad (2.2)$$

i. e. the  $y$  coordinate stays unchanged. This is exactly because the frame  $e^k$  is left-invariant.

As any cotangent space of a smooth manifold, the space  $T^*M$  carries a natural symplectic structure. We start with the canonical 1-form on  $T^*G$ , which is given by

$$\alpha = y_k e^k(a). \quad (2.3)$$

The symplectic form  $\omega$  is then given by

$$\omega = d\alpha. \quad (2.4)$$

We have

$$d\alpha = dy_k \wedge e^k + y_k de^k. \quad (2.5)$$

The calculation of  $de^k$  is standard. First, we introduce the structure constants of  $\mathfrak{g}$  (with respect to the basis  $e_k$  by

$$[e_i, e_j] = c_{ij}^k e_k. \quad (2.6)$$

Next, we apply Cartan's formula for the exterior differentiation:

$$de^k(e_i, e_j) = e_i \cdot e^k(e_j) - e_j \cdot e^k(e_i) - e^k([e_i, e_j]). \quad (2.7)$$

Combining (2.6) and (2.7), together with the fact that the first two terms on the right-hand side of (2.7) vanish due to left-invariance of the objects involved, we obtain

$$\omega = d\alpha = dy_k \wedge e^k - \frac{1}{2} y_k c_{ij}^k e^i \wedge e^j. \quad (2.8)$$

In other words, in the local frame on  $T^*G$  given by  $e_1, \dots, e_n, e^1 \sim \frac{\partial}{\partial y_1}, \dots, e^n \sim \frac{\partial}{\partial y_n}$ , the form  $\omega$  is given by the block matrix <sup>2</sup>

$$\begin{pmatrix} -C(y) & -I \\ I & 0 \end{pmatrix}, \quad (2.9)$$

where  $C(y)$  denotes the matrix  $y_k c_{ij}^k$ . The inverse of the matrix (2.9) is

$$\begin{pmatrix} 0 & I \\ -I & -C(y) \end{pmatrix}, \quad (2.10)$$

and for any function  $H = H(a, y)$  on  $T^*G$  the corresponding Hamiltonian equations are

$$\begin{aligned} (a^{-1}\dot{a})^k &= \frac{\partial H}{\partial y_k}, \\ \dot{y}_k &= -e_k H + y_l c_{jk}^l \frac{\partial H}{\partial y_j}, \end{aligned} \quad (2.11)$$

where  $(a^{-1}\dot{a})^k$  denotes the  $k$ -th coordinate of the vector  $a^{-1}\dot{a} \in \mathfrak{g}$ , the expression  $e_k H$  denotes the derivative along the  $e_k$  direction in the variable  $a$ . The last term on the right-hand side of the second equations represents the Poisson bracket  $\{H, y_k\}$  with  $H$  considered as a functions of  $y$  (and  $a$  considered as fixed when calculating the bracket). The bracket is uniquely given by its usual properties and the relations

$$\{y_i, y_j\} = y_k c_{ij}^k. \quad (2.12)$$

It can be obtained by applying the standard Poisson bracket on the symplectic manifold  $T^*G$  to functions independent of  $a$  in the above coordinates  $(a, y)$ .

Note that the equations (2.11) do not depend on the metric, they depend only on the structure of the Lie algebra.

---

<sup>2</sup>We use the usual identifications: if  $f, g$  are two co-vectors with coordinates  $f_i, g_j$  respectively, then the two-form  $f \wedge g$  is identified with the antisymmetric matrix  $\omega_{ij} = f_i g_j - f_j g_i$  and  $(f \wedge g)(\xi, \eta) = \omega_{ij} \xi^i \eta^j$  for any two vectors  $\xi, \eta$ .

### 2.3 The symplectic reduction to $\mathfrak{g}^*$ and the Euler-Arnold equation

When  $H$  is invariant under the prolongation of the action by left multiplication of  $G$  on itself to  $T^*G$ , which is equivalent to  $H$  not depending on  $a$  in the above coordinates  $(a, y)$ , i. e.  $H = H(y)$ , then the second equation of (2.11) does not contain  $a$  and is simply

$$\dot{y}_k = \{H, y_k\}. \quad (2.13)$$

This is a form of the Euler-Arnold equation, originally formulated in  $\mathfrak{g}$  in [1]. This equation represents one form of the reduction of the equations on  $T^*G$  to  $\mathfrak{g}^*$  by the symmetries of the left action of  $G$  on itself, see for example [15]. The space  $\mathfrak{g}^*$  has a natural structure of a Poisson manifold (with the Poisson bracket given by (2.12)) and is foliated into “symplectic leaves”, which are given by the orbits of the co-adjoint representation, see for example [2, 15]. The orbits are given by

$$\mathcal{O}_{\bar{y}} = \{(\text{Ad } a)^* \bar{y}, a \in G\} \quad (2.14)$$

where  $\bar{y}$  is a fixed vector in  $\mathfrak{g}^*$  and  $\text{Ad } a$  is defined below, and they have a natural structure of a symplectic manifold (with the maps  $(\text{Ad } a)^*$  acting by as symplectic diffeomorphism).

### 2.4 Conserved quantities, the moment map, and Noether’s theorem

The Killing fields associated with the symmetries of the Riemannian structure on  $G$  with the left-invariant metric given by left multiplications  $b \rightarrow ab$  are easily seen to be given by *right-invariant* vector fields  $e(a) = \xi a$  (where  $\xi \in \mathfrak{g}$ ) on  $G$ . By Noether’s theorem there should be a conserved quantity associated to any such field. It is easy to see that the quantity is given by

$$(a, y) \rightarrow ((\text{Ad } a^{-1}) \xi, y) = (\xi, (\text{Ad } a^{-1})^* y), \quad (2.15)$$

where the operator  $\text{Ad } a$  is defined as usual by

$$\text{Ad } a \xi = a \xi a^{-1}. \quad (2.16)$$

The map  $M: T^*G \rightarrow \mathfrak{g}^*$  given in the  $(a, y)$  coordinates by

$$M(a, y) = (\text{Ad } a^{-1})^* y \quad (2.17)$$

is the usual moment map associated with the (symplectic) action of  $G$  on  $T^*G$  (given by the prolongation of the left multiplication). The vector  $M(a, y)$  is conserved under the Hamiltonian evolution, and the quantities  $(\xi, M)$  are the conserved quantities from Noether’s theorem applied to our situation. In particular, the Hamiltonian equations (2.11) obtained from taking the Hamiltonian as

$$H(a, y) = (\xi, M(a, y)) \quad (2.18)$$

are

$$\begin{aligned} (\dot{a})a^{-1} &= \xi \\ \dot{y} &= 0. \end{aligned} \quad (2.19)$$

The conservation of  $M$  also has a geometric interpretation: if  $x(s)$  is a geodesics (parametrized by length) on a Riemannian manifold and  $X$  is a Killing field (infinitesimal symmetry), then the scalar product  $(\dot{x}, X)$  is constant. This is of course just another way to state the Noether’s theorem in this particular case, but it can also be interpreted in terms of properties of Jacobi fields along our geodesics.

In the context of rotating rigid bodies, the quantity  $M$  corresponds to the conservation of angular momentum, see [1]. In the context of ideal fluids, the conservation of  $M$  corresponds to the Kelvin-Helmholtz laws for vorticity, as observed by many authors.

It is easy to check the following fact: when  $H$  is independent of  $a$ , i. e.  $H = H(y)$ , then for a curve  $(a(t), y(t))$  in  $T^*G$  satisfying the “kinematic” equation

$$(a^{-1}\dot{a})^k = \frac{\partial H}{\partial y_k} \quad (2.20)$$

the “dynamical” equation

$$\dot{y}_k = \{H, y_k\} \quad (2.21)$$

is equivalent to the (generalized) momentum conservation

$$M(a, y) = \text{const.} \quad (2.22)$$

Also, if  $(a(t), y(t))$  is a solution of the equations of motion and  $H = H(y)$ , then  $y(t)$  is given by

$$y(t) = (\text{Ad } a(t))^* \bar{y} \quad (2.23)$$

for some fixed co-vector  $\bar{y} \in \mathfrak{g}^*$ .

### 3 Perturbations by random forces

#### 3.1 Langevin equation

A very natural random perturbation of the geodesic flow is the Langevin equation, which can be symbolically written as

$$\ddot{a} = -\nu\dot{a} + \varepsilon\dot{w}, \quad (3.1)$$

for some parameters  $\nu > 0$  and  $\varepsilon > 0$ , which for a given  $t > 0$  and  $a(t)$  is considered as an equation in  $T_{a(t)}G$ , with  $\ddot{a}$  interpreted as the covariant derivative of  $\dot{a}$  along the curve  $a(t)$ , and  $w$  is a suitable form of Brownian motion in the Riemannian manifold  $G$ . Of course, the expression  $\dot{w}$  is somewhat ambiguous and there are some subtle points in writing things in the correct way from the point of view of rigorous stochastic calculus. In particular, one has to distinguish carefully between the Itô and Stratonovich integrals. Here we will mostly avoid the subtleties of the right interpretation of the stochastic equations such as (3.1) by working instead with the Fokker-Planck equation, and we can define the transition probabilities for our processes via that equation.

A good starting point for writing the Fokker-Planck equation associated with (3.1) is the Liouville equation in  $T^*G$ . This equation describes the evolution of a density  $f(a, y)$  (with respect to the volume element given by the natural extension of the Riemannian metric from  $G$  to  $T^*G$ , which is proportional to the volume element given by the  $n$ -th power  $\omega \wedge \omega \wedge \cdots \wedge \omega$  ( $n$  times) of the canonical symplectic form  $\omega$  above. The Liouville equation is

$$f_t + v^k e_k f + b_k \frac{\partial f}{\partial y_k} = 0, \quad (3.2)$$

where

$$v^k = \frac{\partial H}{\partial y_k}, \quad b_k = \{H, y_k\}, \quad (3.3)$$

and  $e_k f$  denotes the differentiation of  $f(a, y)$  as a function of  $a$  in the direction of the field  $e_k$  defined earlier. The vector field  $X = v^k e_k + b_k \frac{\partial}{\partial y_k}$  is div-free (with respect to our volume form

in  $T^*G$ ), as follows from the Liouville theorem in Hamiltonian mechanics. Hence equation (3.2) is the same as

$$f_t + \operatorname{div}(Xf) = 0, \quad (3.4)$$

where  $\operatorname{div}$  is taken in our metric on  $T^*G$ . Our Fokker-Planck equation should then be

$$f_t + v^k e_k f + b_k \frac{\partial f}{\partial y_k} + \frac{\partial}{\partial y_k} \left( -\nu v_k f - \frac{\varepsilon^2}{2} \frac{\partial f}{\partial y^k} \right) = 0, \quad (3.5)$$

where  $v_k$  is as above. It can be considered as a combination of the Liouville transport with an Ornstein-Uhlenbeck process along the linear fibers of  $T^*G$ . To have an exact correspondence to (3.1), we have to take  $H(y) = \frac{1}{2}|y|^2$ , the Hamiltonian of the geodesic flow. This equation can then be interpreted as describing a “physical Brownian motion” of a particle in  $G$ . (We can for example think of  $G$  being filled with an incompressible fluid which is at rest and the Brownian particle suspended in the fluid and being subject to random “kicks” from the fluid molecules and friction due to viscosity of the fluid, in the spirit of Einstein’s paper [4].

The symmetry reduction for (3.5) corresponding to the symmetry reduction for (3.1) is very simple: we consider it only for functions depending on  $y$ , which results in dropping the term  $v^k e_k f$ . The symplectic reduction of (2.11) to (2.13) corresponds to the same procedure applied to the Liouville equation (3.2).

There is an explicit steady solution of (3.5) given by

$$f(a, y) = C e^{-\beta H}, \quad \beta = \frac{2\nu}{\varepsilon^2}, \quad (3.6)$$

where  $C$  is any constant. The formula is the same as in the flat space. The approach to equilibrium will however be influenced by the term  $b_k \frac{\partial}{\partial y_k}$  which is absent in the flat case. Strictly speaking, the last statement applies unambiguously only to compact groups  $G$ , where the equilibrium (3.6) is easily seen to be unique among probability densities (for a suitable  $C$ , under some natural assumptions on  $H$ ). We will discuss this point in some detail below in the more difficult case of degenerate forcing.

Given that the conservation of  $M(a, y)$ , from the point of view of Statistical Mechanics it is natural to consider (at least when  $G$  is compact) distributions in the phase space  $T^*G$  given by

$$f(a, y) = C e^{-\beta H(y) + (\xi, M(a, y))} = C e^{-\beta H(y) + ((\operatorname{Ad} a^{-1})\xi, y)} \quad (3.7)$$

for  $\beta > 0$  and  $\xi \in \mathfrak{g}$ . In fact, if we replaced the Langevin equation by the Boltzmann equation

$$f_t + v^k e_k f_k + b_k \frac{\partial f}{\partial y_k} = Q(f, f), \quad (3.8)$$

for appropriate “collision operator”  $Q$  (defined on each fiber  $T_a^*G$  in the same way as in the flat case), densities (3.7) should be among the equilibria (the set of which could possibly be larger due to symmetries other than those generated by the left shifts). The large degeneracy of the set of equilibria is an important feature of the Boltzmann equation which is crucial for fluid mechanics. It is not shared by the Langevin equation, for which the equilibrium is unique (under reasonable assumptions). This is related to the hypoellipticity of the differential operator in (3.5), which we will discuss in some detail for more general operators in the next subsection.

We remark that one can modify the Langevin equation and get (3.7) as equilibria for the modified equation. For this we simply change the Hamiltonian in (3.5) to the expression

$$\tilde{H}(a, y) = H(y) - ((\operatorname{Ad} a^{-1})\xi, y) \quad (3.9)$$

This corresponds to watching a Brownian motion of a particle in incompressible fluid which moves in  $G$  as a rigid body along the Killing field  $\xi a$ . (This is a steady solution of the equations of motion for an incompressible fluid.) The term  $((\text{Ad } a^{-1})\xi, y)$  in the Hamiltonian is then produces the analogues of centrifugal and Coriolis forces which we encounter in rotating coordinate frames.

### 3.2 Langevin equation with degenerate forcing

In PDEs of fluid mechanics one sometimes considers forcing through low spatial Fourier modes which is “white noise” in time. See for example [8, 14]. In our context here this is akin to considering the system

$$\begin{aligned} (a^{-1}\dot{a})^k &= \frac{\partial H}{\partial y_k} , \\ \dot{y}_k &= -c_k H + y_l c_{jk}^l \frac{\partial H}{\partial y_j} - \nu \frac{\partial H}{\partial y_k} + \sum_{i=1}^r \varepsilon \dot{w}_i \tilde{f}_k^i , \end{aligned} \quad (3.10)$$

where  $\tilde{f}^1, \dots, \tilde{f}^r$  are some fixed vectors in  $\mathfrak{g}^*$  and  $w_i$  are standard independent Wiener processes. The term  $-\nu \frac{\partial H}{\partial y_k}$  represents friction. One could consider more general forms of frictions, but here we will be content with the above special form. Many of the results below hold for more general friction forces. The main complication in (3.10) as compared to the previous section is that  $r$  can be less than the dimension  $n$  of  $\mathfrak{g}^*$ .

In the remainder of this subsection we will assume that

$$H = H(y) = \frac{1}{2}|y|^2 = \frac{1}{2}y_k y^k , \quad (3.11)$$

which correspond to geodesic flow, or kinetic energy in classical mechanics. Also, below we will need to do some Lie bracket calculations for which some formulae seem to be easier when we work in  $\mathfrak{g}$  rather than  $\mathfrak{g}^*$ . This amounts to “raising the indices” in the old-fashioned language, i. e. working in the coordinates  $y^k$  rather than  $y_k$ . We note that with these assumptions we have

$$y^k = v^k . \quad (3.12)$$

Equation (3.10) then becomes

$$\begin{aligned} (a^{-1}\dot{a})^k &= y^k , \\ \dot{y}^k &= \tilde{q}_{ij}^k y^i y^j - \nu y^k + \sum_{i=1}^r \varepsilon \dot{w}^i f_i^k , \end{aligned} \quad (3.13)$$

where the notation is self-explanatory, perhaps with the exception of the term  $\tilde{q}_{ij}^k y^i y^j$ , in which the coefficients are not uniquely determined by the function  $y \rightarrow \tilde{q}(y, y)$ . A straightforward “raising of indices” gives the definition

$$([x, y], z) = (\tilde{q}(z, x), y) , \quad x, y, z \in \mathfrak{g} , \quad (3.14)$$

which coincides with the Arnold form  $B$  from [1]. In what follows it will be advantageous to work with the symmetrization of  $\tilde{q}$ , which will be denoted by  $q$ :

$$q(x, y) = \frac{1}{2}(\tilde{q}(x, y) + \tilde{q}(y, x)) . \quad (3.15)$$

In equation (3.13) it does not matter whether we use  $\tilde{q}$  or  $q$ , of course. Instead of (3.13) we can write

$$\begin{aligned} \dot{a} &= az , \\ \dot{z} &= q(z, z) - \nu z + \varepsilon \sigma \dot{w} , \end{aligned} \quad (3.16)$$



where we use  $z$  to emphasize that the equations are considered in  $\mathfrak{g}$ , as the variable  $y$  was used to denote elements of  $\mathfrak{g}^*$ ,  $w$  is the vector of the standard Wiener process in  $R^r$  and  $\sigma$  is a suitable  $n \times r$  matrix. The corresponding Fokker-Planck equation for  $f = f(a, z; t)$  then is

$$f_t + z^k e_k f + q^k(z, z) \frac{\partial f}{\partial z^k} + \frac{\partial}{\partial z^k} \left( -\nu z^k f - \frac{\varepsilon^2}{2} h^{kl} \frac{\partial f}{\partial z^l} \right), \quad (3.17)$$

for a suitable symmetric positive semi-definite matrix  $h$  (which is constant in  $z$ ).

This is clearly a degenerate parabolic operator and we will study the classical parabolic Hörmander condition for hypoellipticity for the Lie brackets generated by the vector fields relevant for the operator, see [6]. In our context here the condition can be formulated in terms of the “Lie algebra hull” of the vector fields

$$\mathcal{X}_k = \sigma_k^l \frac{\partial}{\partial z^l} \quad (3.18)$$

which satisfies the crucial additional condition that it is closed under the operation

$$\text{Ad } \mathcal{X}_0 : \mathcal{X} \rightarrow [\mathcal{X}_0, \mathcal{X}], \quad (3.19)$$

where the last bracket is the Lie bracket of vector fields and  $\mathcal{X}_0$  is the “Euler-Arnold component” of our operator as specified below.

The coordinates on  $TG$  we will use are  $(a, z)$ , which correspond to  $z^k e_k(a)$ . The vector fields on  $TG$  which will be relevant for our purposes will be of the form  $A^k(z) e_k(a) + X^k(z) \frac{\partial}{\partial z^k}$ . We will write

$$A^k(z) e_k(a) + X^k(z) \frac{\partial}{\partial z^k} = \begin{pmatrix} A \\ X \end{pmatrix} = \begin{pmatrix} A(z) \\ X(z) \end{pmatrix}. \quad (3.20)$$

In these coordinates the Lie bracket is

$$\left[ \begin{pmatrix} A \\ X \end{pmatrix}, \begin{pmatrix} B \\ Y \end{pmatrix} \right] = \begin{pmatrix} A \wedge B + D_X B - D_Y A \\ [X, Y] \end{pmatrix}, \quad (3.21)$$

where we use  $A \wedge B$  to denote the function of  $z$  obtained from  $A(z)$  and  $B(z)$  by taking the Lie bracket in  $\mathfrak{g}$  pointwise, as opposed to  $[X, Y]$ , which denotes the Lie bracket of  $X, Y$  considered as vector fields in  $\mathfrak{g}$ . The notation  $D_A X$  has the usual meaning: the derivative of  $X = X(z)$  (at  $z$ ) in the direction of  $A = A(z)$ .

Let us write  $Q = Q(z, z)$  for the vector field in  $\mathfrak{g}$  given by the vector field  $q^k(z, z) \frac{\partial}{\partial z^k}$ .

For simplicity we will work out the case when  $h^{kl}$  is of rank one, which means that the random forcing is applied only in one direction, which will be denoted by  $F$  (and considered as a constant vector field in  $\mathfrak{g}$ ). Hence

$$h^{kl} = F^k F^l. \quad (3.22)$$

In this case the vector fields for the Hörmander condition calculation can be taken as

$$\begin{pmatrix} 0 \\ F \end{pmatrix}, \quad \text{and} \quad \mathcal{X}_0 = \frac{1}{2} \begin{pmatrix} z \\ Q - \nu z \end{pmatrix}, \quad (3.23)$$

We have

$$\left[ \begin{pmatrix} 0 \\ F \end{pmatrix}, \begin{pmatrix} z \\ Q - \nu z \end{pmatrix} \right] = \begin{pmatrix} F \\ D_F Q - \nu F \end{pmatrix} \quad (3.24)$$

and

$$\left[ \begin{pmatrix} 0 \\ F \end{pmatrix}, \left[ \begin{pmatrix} 0 \\ F \end{pmatrix}, \begin{pmatrix} z \\ Q - \nu z \end{pmatrix} \right] \right] = \begin{pmatrix} 0 \\ D_F^2 Q \end{pmatrix}. \quad (3.25)$$

This means that we have extended our list of vector fields by the field

$$\begin{pmatrix} 0 \\ G \end{pmatrix}, \quad G = \frac{1}{2}D^2Q = Q(F, F). \quad (3.26)$$

We can now take

$$\left[ \begin{pmatrix} 0 \\ G \end{pmatrix}, \left[ \begin{pmatrix} 0 \\ F \end{pmatrix}, \begin{pmatrix} z \\ Q - \nu z \end{pmatrix} \right] \right] = \begin{pmatrix} 0 \\ D_G D_F Q \end{pmatrix}. \quad (3.27)$$

and extend our list of fields by

$$\begin{pmatrix} 0 \\ Q(F, G) \end{pmatrix}. \quad (3.28)$$

We note that the new fields obtained in this way are “constant” (in the coordinates we use), so the procedure can be easily iterated.

**Definition 3.1** *We will say that  $Q$  is non-degenerate with respect to a set  $\mathcal{F} \subset \mathfrak{g}$  if there is no non-trivial subspace  $M \subset \mathfrak{g}$  containing  $\mathcal{F}$  which is invariant under  $Q$ , in the sense that  $Q(z, z') \in M$  whenever  $z, z' \in M$ .*

We can now formulate the main result of this subsection:

**Theorem 3.1** *The operator of the Fokker-Planck equation (3.17) satisfies the Hörmander condition if and only if  $Q$  is non-degenerate with respect to the range of the matrix  $h$  (considered as a map from  $\mathfrak{g}$  to  $\mathfrak{g}$ ).*

*Proof:* The necessity of the condition can be seen when we consider functions depending only on  $z$ . If there is a non-trivial linear subspace invariant under both  $Q$  and the diffusion, then particle trajectories starting at  $M$  clearly cannot leave  $M$ , and therefore the operator cannot satisfy the Hörmander condition.

On the other hand, if  $Q$  is non-degenerate with respect to the range of  $h$ , then the above calculation shows that the Lie brackets of the fields (3.23) (with perhaps several fields of the same form as the first one) generate the fields of the form

$$\begin{pmatrix} 0 \\ X_j \end{pmatrix}, \quad j = 1, \dots, n, \quad (3.29)$$

where  $X_1, \dots, X_n \in \mathfrak{g}$  form a basis of  $\mathfrak{g}$ . Formula (3.23) now shows that the fields of the form

$$\begin{pmatrix} X_j \\ Y_j(z) \end{pmatrix} \quad (3.30)$$

can also be generated. Together with the fields (3.29) they clearly form a basis of  $T(TG)$  at each point  $(a, z)$ , and the proof is finished.

*Remarks* 1. If one replaces the damping term  $-\nu z$  in (3.16) by a more general expression  $-\nu Dz$ , where  $D$  is a positive-definite matrix, interesting new questions arise. We plan to address these in a future work.

2. Very recently we learned about the paper [10]. The results there could be used (modulo simple modifications) to prove the above theorem and also to say more about the set where the solutions of the Fokker-Planck equation are positive.

**Corollary 3.1** *When  $G$  is compact and  $Q$  is non-degenerate with respect to the diffusion matrix  $h$  in the sense above, the process (3.16) (and hence also (3.13)) is ergodic with respect to a distribution density given by a function which is independent of  $a$ . In other words, the Lagrangian positions of the “particles” are uniformly distributed (with respect to the Haar measure) in the limit of infinite time.*

In our situation this is not hard to prove once the Hörmander condition is established by following methods in [7, 6] and [13].

*Remark.* One should be also able to prove convergence to the equilibrium measure following the methods of [17], but we will not pursue this direction here. It is perhaps worth reminding that in general there is a difference between uniqueness of the ergodic measure and the convergence to equilibrium. A simple example in our context here is provided by the equation

$$f_t + f_{x_1} = \frac{1}{2} f_{x_2 x_2} . \quad (3.31)$$

considered in the 2d torus. Note that this equation does not satisfy the parabolic Hörmander condition, while its spatial part satisfies the elliptic Hörmander condition.

### 3.3 Constrained diffusion in the momentum space

The Euler-Arnold equation (2.13) leaves invariant the co-adjoint orbits (2.14) and also the energy levels  $\{H = E\}$ . It is therefore of interest to consider perturbations by noise which “respects” some of the constraints. For example, one can add noise respecting the coadjoint orbit, but not the energy levels. An example of this situation (in the presence on non-holonomic constraints) is considered in [11]. It is closely related to stochastic processes on co-adjoint orbits introduced by Bismut [3]. One can also consider noise which preserve energy levels but not necessarily the co-adjoint orbits, or one can consider noise which preserves both the coadjoint orbits and the energy levels.

We wish to include all these situations in our considerations, and therefore we will consider the following scenario. We assume that we are given a Hamiltonian  $H = H(y)$  and a manifold  $M \subset \mathfrak{g}^*$  which is invariant under the evolution by the Euler-Arnold equation (2.13). If  $M$  is given locally as a non-degenerate level set of some conserved quantities  $\phi_1, \dots, \phi_r$  (in the sense that  $\{H, \phi_k\} = 0$ ,  $k = 1, \dots, r$ ),

$$M = \{y \in \mathfrak{g}^* , \phi_1(y) = c_1, \phi_2(y) = c_2, \dots, \phi_r(y) = c_r\} \quad \text{locally in } y, \quad (3.32)$$

there is a natural measure  $m$  on  $M$  (invariant for the Hamiltonian flow) which is generated by the volume in  $\mathfrak{g}^*$  (given by our metric there) and the conserved quantities by first restricting the volume measure in  $\mathfrak{g}^*$  to

$$M_\varepsilon = \{y \in \mathfrak{g}^* , \phi_1(y) \in (c_1 - \varepsilon, c_1 + \varepsilon), \phi_2(y) \in (c_2 - \varepsilon, c_2 + \varepsilon), \dots, \phi_r(y) \in (c_r - \varepsilon, c_r + \varepsilon)\} \quad (3.33)$$

then normalizing the restricted measure by a factor  $\frac{1}{2\varepsilon}$  and taking the limit  $\varepsilon \rightarrow 0_+$ . In the case  $r = 1$  we have

$$m = \frac{1}{|\nabla \phi_1|} \mathcal{H}^{n-1}|_M , \quad (3.34)$$

where  $\mathcal{H}^{n-1}$  is the  $n - 1$  dimensional Hausdorff measure generated by our metric, and the gradient and its norm in the formula are also calculated with our given metric. For general  $r$  we have similar formulae, the corresponding expression can be seen easily from the co-area

formula, for example. However, the above definition via the limit  $\varepsilon \rightarrow 0_+$  is perhaps more natural, as it relies only on the objects which are “intrinsic” from the point of view of the definition of  $m$ : the underlying measure in  $\mathfrak{g}^*$  and the constraints  $\phi_k$ . (The proof that the limit as  $\varepsilon \rightarrow 0_+$  is well-defined is standard and is left to the interested reader.)

As the Hamiltonian evolution in the phase-space  $T^*G \sim G \times \mathfrak{g}^*$  preserves the Liouville measure, which, in the  $(a, y)$  coordinates defined by (2.1), is the product of the Haar measure on  $G$  and the canonical volume measure in  $\mathfrak{g}^*$ , we see that the product of the (left) Haar measure  $h_G$  on  $G$  and  $m$  is an invariant measure for the Hamiltonian evolution in the subset of  $T^*G$  given by  $G \times M$  in the  $(a, y)$  coordinates. If the group  $G$  is not unimodular<sup>3</sup>, the measure  $m$  may not be preserved by the Euler-Arnold equation (2.13) in  $\mathfrak{g}^*$ , which represents the symplectic reduction of the original full system. This is because while the vector field

$$v^k e_k + q_k \frac{\partial}{\partial y_k} \quad (3.35)$$

in the Liouville equation (3.2) is div-free in  $T^*G$ , its two parts may not be div-free in  $G$  or  $\mathfrak{g}^*$  respectively, unless the group is unimodular.

The Liouville equation for the evolution in  $G \times M$  is the same as (3.2)

$$f_t + v^k e_k f + b_k \frac{\partial f}{\partial y_k} = 0, \quad (3.36)$$

where  $f = f(a, y)$  now denotes the density with respect to the measure  $h_G \times m$  (where  $h_G$  is again the left Haar measure on  $G$ ).

We now consider stochastic perturbations of the Liouville equation (3.36) on  $G \times M$ . As in the Langevin-type equations, the random forces will act only in the  $y$ -component, so that the kinematic equation  $(a^{-1}\dot{a})^k = v^k$  is left unchanged.

We will demand that the stochastic term will also leave invariant the measure  $h_G \times m$ , and as it acts only in the  $y$ -variable, it then must leave invariant the measure  $m$ .

There is more than one way in which noise can be introduced in a reasonable way into (3.36). For example, if  $V$  is a vector field (with coordinates  $V_k$ ) tangent to  $M$  which generates a flux on  $M$  preserving the measure  $m$ , one can replace the equation (2.11) by

$$dy_k = \{H, y_k\} dt + \varepsilon V_k \circ dW, \quad (3.37)$$

where  $W$  is the standard 1d Wiener process and  $\circ$  indicates, as usual, that the corresponding stochastic integrals should be taken in the Stratonovich sense.<sup>4</sup> The corresponding Fokker-Planck equation is given by

$$f_t + v^k e_k f + b_k \frac{\partial f}{\partial y_k} = \frac{\varepsilon^2}{2} (D_V^*)^2 f, \quad (3.38)$$

where  $D_V^*$  is adjoint to  $D_V = V_k \frac{\partial}{\partial y_k}$  with respect to the measure  $m$ . As the flux by  $V$  preserves  $m$ , we in fact have  $D_V^* = -D_V$ . In this case the operators  $D_V^2$  and  $(D_V^*)^2$  coincide and arise from the functional

$$\int_M \frac{1}{2} |D_V f|^2 dm \quad (3.39)$$

This is in some sense the “minimal non-trivial noise” model, and it might be of interest in some situations.

---

<sup>3</sup>Recall that a group is unimodular if the notions of left invariant and right invariant Haar measures coincide. This is the same as demanding that the maps  $y \rightarrow \text{Ad } a^* y$  preserve the volume in  $\mathfrak{g}^*$ , i. e. have determinant 1.

<sup>4</sup>Note that with Itô integration the particle trajectories might not stay in  $M$ .

Here we will consider the situation when the noise is non-degenerate in  $M$ , leaving the interesting case of the degenerate noise in  $M$  to future studies. Our motivation is the following. For  $\varepsilon > 0$  we consider the usual Brownian motion in  $\mathfrak{g}^*$ , but restricted to the set  $M_\varepsilon$  above, with the understanding that the trajectories “reflect back” (we can think about an action of some control mechanism) at the boundary (corresponding to the Neumann condition at the boundary for the corresponding Fokker-Planck equation, which is just the heat equation in this case). As  $\varepsilon \rightarrow 0_+$ , a good model for the limiting process on  $M$  is given by an operator constructed as follows. First, we take the metric induced on  $M$  by the given metric in  $\mathfrak{g}^*$ . Assume the metric is given by  $\tilde{g}_{ij}$  in some local coordinates. Assume the measure  $m$  is given as  $m(x) dx$  in these coordinates, where  $m(x)$  is a (smooth) function. Denoting by  $\tilde{g}$  the determinant of  $\tilde{g}_{ij}$ , the volume element given by  $\tilde{g}_{ij}$  in our coordinates is  $\sqrt{\tilde{g}} dx$ . We then define a new metric

$$h_{ij} = \varkappa \tilde{g}_{ij} \quad (3.40)$$

so that the volume element  $\sqrt{h} dx$  satisfies

$$\sqrt{h} dx = m(x) dx . \quad (3.41)$$

Then we take the generator of our process to be the Laplace operator on  $M$  with respect to the metric  $h$ . We will denote this operator by  $L_M$ . Our Fokker-Planck equation then will be

$$f_t + v^k e_k f + b_k \frac{\partial f}{\partial y_k} = \frac{\varepsilon^2}{2} L_M f . \quad (3.42)$$

We will be interested in ergodicity properties of the process given by this equation.

In the remainder of this subsection we will assume again (3.11), i. e. the Hamiltonian  $H$  is quadratic (and positive definite). We can then “lower the indices” and work with  $TG$  and  $\mathfrak{g}$  rather than with  $T^*G$  and  $\mathfrak{g}^*$ . We will denote by  $Z$  the image of  $M$  in  $\mathfrak{g}$  under the “lowering indices” map, and will denote the elements of  $Z \subset \mathfrak{g}$  by  $z$ , with coordinates  $z^k$ . Similarly to (3.12) we have  $z^k = v^k$ . The Fokker-Planck equation (3.42), now considered on  $G \times Z$  becomes

$$f_t + z^k e_k f + q^k(z, z) \frac{\partial f}{\partial z^k} = \frac{\varepsilon^2}{2} L f , \quad (3.43)$$

where  $q^k$  is defined by (3.15), and  $L$  is the operator on  $Z$  corresponding to  $L_M$ . It is of course again a Laplacian for some metric on  $Z$  (which is conformally equivalent to the metric on  $Z$  induced by the underlying metric in  $\mathfrak{g}$ ).

Let us now consider conditions under which the operators corresponding to (3.43) or (3.42) satisfy the usual Hörmander commutator condition for hypoellipticity.

**Definition 3.2** *A  $p$ -hull<sup>5</sup> of a subset  $S \subset \mathfrak{g}$  is the smallest Lie sub-algebra  $\mathfrak{h} \subset \mathfrak{g}$  with the following properties:*

- (i)  $\mathfrak{h}$  contains the set  $S - S = \{s_1 - s_2, , s_1, s_2 \in S\}$ ,
- (ii)  $\mathfrak{h}$  is invariant under the mappings  $\text{Ad } s: z \rightarrow [s, z]$  for each  $s \in S$ .

---

<sup>5</sup>Here  $p$  stands for parabolic, as the definition is tied to the parabolic Hörmander condition.

Remarks:

1. The p-hull will be relevant in the context of the evolution equation (3.43). For the “spatial part” of the operator (3.43), obtained by omitting the term  $f_t$ , the relevant “hull” is simply the Lie algebra generated by  $S$ .

2. Condition (i) in the definition already implies that  $\mathfrak{h}$  is invariant under the mapping  $\text{Ad}(s_1 - s_2)$  for any  $s_1, s_2 \in S$ . Therefore in (ii) we can require invariance of  $\mathfrak{h}$  under  $\text{Ad } s_0$  for just one fixed  $s_0 \in S$  (and - given (i) - the definition will be independent of the choice of  $s_0$ ).

The main result of this section is the following:

**Theorem 3.2** *In the notation introduced above, assume that  $M$  is a smooth analytic submanifold of  $\mathfrak{g}^*$ . Then the operator on  $G \times M$  corresponding to (3.42) (or, equivalently, the operator on  $G \times Z$  corresponding to (3.43)) satisfies the Hörmander condition if and only if the p-hull of  $Z$  coincides with  $\mathfrak{g}$ .*

*Proof:*

Let us first show that the p-hull condition is necessary for the Hörmander condition. One can see this from the Lie bracket calculations below, but it is instructive to verify it directly. Assume  $\mathfrak{h}$  is a non-trivial Lie subalgebra of  $\mathfrak{g}$  containing  $Z - Z$  for which we can find  $z_0 \in Z \setminus \mathfrak{h}$  such that  $\mathfrak{h}$  is invariant under  $\text{Ad } z_0$ . Let us set

$$e = z_0^k e_k . \quad (3.44)$$

The Lie algebra  $\mathfrak{h}$  defines (locally) a foliation  $\mathcal{F}$  of  $G$  into cosets  $aH$ , where  $H$  is the (local) Lie subgroup of  $G$  corresponding to  $\mathfrak{h}$ . The main point now is that the invariance of  $\mathfrak{h}$  under  $\text{Ad } z_0$  implies that the flow given by the equation

$$a^{-1} \dot{a} = e \quad (3.45)$$

preserves the foliation. (Another formulation of this statement could be that the equation (3.45) “descends” to  $G/H$ .) This means that the perturbations given by the stochastic terms in (3.43) will still preserve the foliation (e. g. by the Stroock-Varadhan Theorem [16]) and it is not hard to conclude that set of points reachable by the corresponding process cannot be open.

For the proof that the p-hull condition is sufficient, we write our operator (locally) in the form

$$f_t + \mathcal{X}_0 f - \sum_{j=1}^m \mathcal{X}_j^2 f , \quad (3.46)$$

where  $m$  is the dimension of  $Z$  (which is of course the same as the dimension of  $M$ ) and  $\mathcal{X}_j$  are suitable vector fields on  $G \times Z$ . All these fields will be of the form (3.20), and we will use the same notation as in (3.20) in what follows. We will be working locally near a point  $(a, z_0) \in G \times Z$ . We choose  $\mathcal{X}_j, j = 1, \dots, m$  so that

$$\mathcal{X}_j = \begin{pmatrix} 0 \\ Y_j \end{pmatrix} , \quad j = 1, \dots, m . \quad (3.47)$$

where  $Y_j$  are analytic near  $z_0$  and  $Y_j(z)$  form a basis of  $T_z Z$  for each  $z$  close to  $z_0$ . The field  $\mathcal{X}_0$  will be of the form

$$\mathcal{X}_0 = \begin{pmatrix} z \\ V \end{pmatrix} , \quad (3.48)$$

where  $V$  is an field on  $Z$  (analytic near  $z_0$ ). Let us consider local analytic vector fields on  $G \times Z$  near  $(a, z_0)$  of the form

$$\mathcal{X}(a, z) = \begin{pmatrix} X(z) \\ Y(z) \end{pmatrix} \quad (3.49)$$

as a module  $\mathcal{A}$  over the set of analytic functions of  $z \in Z$ . (Recall that we assume that  $Z$  is analytic.) Let  $\mathcal{M}$  be the minimal submodule of  $\mathcal{A}$  satisfying the following requirements:

- (a)  $\mathcal{M}$  contains  $\mathcal{X}_1, \mathcal{X}_2, \dots, \mathcal{X}_m$ , and
- (b)  $\mathcal{M}$  is invariant under the map  $\text{Ad } \mathcal{X}_0: \mathcal{X} \rightarrow [\mathcal{X}_0, \mathcal{X}]$ , where  $[\cdot, \cdot]$  denotes the Lie bracket for vector fields.

The parabolic Hörmander condition at  $(a, z_0)$  for the fields  $\mathcal{X}_0, \mathcal{X}_1, \dots, \mathcal{X}_m$  then is that

$$\{\mathcal{X}(a, z_0), \mathcal{X} \in \mathcal{M}\} = T_{(a, z_0)}(G \times Z). \quad (3.50)$$

For  $\mathcal{X} \in \mathcal{A}$  we will denote by  $\pi\mathcal{X} \in \mathfrak{g}$  the projection to the first component in the notation (3.20), i. e.

$$\pi \begin{pmatrix} X \\ Y \end{pmatrix} = X. \quad (3.51)$$

Let

$$M = \pi\mathcal{M}. \quad (3.52)$$

As  $\mathcal{M}$  contains the vector fields (3.47), the condition (3.50) is equivalent to

$$M_{z_0} = \{X(z_0), X \in M\} = \mathfrak{g}. \quad (3.53)$$

Using (3.21) and the fact that the fields  $\mathcal{X}_1, \dots, \mathcal{X}_m$  belong to  $\mathcal{M}$ , it is we see that  $M$  has the following properties.

$$\text{If } Y \text{ is an analytic vector field on } Z \text{ (defined locally near } z_0), \text{ then } Y \in M. \quad (3.54)$$

This follows by taking the Lie bracket of  $\begin{pmatrix} 0 \\ Y \end{pmatrix}$  and  $\mathcal{X}_0$ .

$$\text{If } A \in M \text{ and } Y \text{ is an analytic vector field on } Z \text{ (defined locally near } z_0), \text{ then } D_Y A \text{ is in } M. \quad (3.55)$$

This follows by taking the Lie bracket of  $\begin{pmatrix} 0 \\ Y \end{pmatrix}$  and an  $\mathcal{X}$  with  $\pi\mathcal{X} = A$ .

$$\text{If } A, B \in M, \text{ then } A \wedge B \text{ is in } M. \quad (3.56)$$

This follows by taking the Lie bracket of the fields  $\mathcal{X}$  and  $\mathcal{Y}$  with  $\pi\mathcal{X} = A$  and  $\pi\mathcal{Y} = B$  and then using (3.55).

$$\text{If } A \in M, \text{ then } z \wedge A \text{ is in } M. \quad (3.57)$$

This follows by taking the Lie bracket of  $\mathcal{X}$  with  $\pi\mathcal{X} = A$  with  $\mathcal{X}_0$  and using (3.54), (3.55), and (3.56). Taking these properties of  $M$  into account, it is clear that the proof of the theorem will be finished if we show that

$$Z - Z \subset M_{z_0}. \quad (3.58)$$

Let  $l$  be a linear function in  $\mathfrak{g}$  which vanishes on  $M_{z_0}$ . As  $Z$  is analytic, the function  $l$  considered as a function on the manifold  $Z$  will be analytic. The property (3.55) of  $M$  implies that the derivatives of all orders  $\geq 1$  of  $l$  at  $z_0$  vanish, and therefore  $l$  must be constant on  $Z$ . In particular,  $l$  must vanish on  $Z - Z$ . We see that no point of  $Z - Z$  can be separated from the subspace  $M_{z_0}$  by a linear function, and (3.58) follows. This finishes the proof of the theorem.

**Corollary 3.2** *If the assumptions of Theorem 3.2 are satisfied and the group  $G$  is compact, then any solution of the Fokker-Planck equation (3.42) approaches constant. In particular, the system is ergodic for the (stochastic) dynamics, with the unique ergodic measure given by the constant density  $f$ .*

*Proof:* We note that

$$\frac{d}{dt} \int_{G \times Z} f^2(a, z, t) da m(dz) = -\varepsilon^2 \int_{G \times Z} |\nabla_z f(a, z, t)|^2 da m(dz), \quad (3.59)$$

where we take  $Z$  with the metric defining  $L$ . By regularity which follows from the Hörmander condition we can consider the  $\Omega$ -limit set  $\Omega(f_0)$  of the evolution starting with  $f_0$ , and it consists of smooth functions. Moreover the integral on the right of (3.59) has to vanish identically for each function in  $\Omega(f_0)$ , by the usual Lyapunov-function-type arguments. This means that any function in  $\Omega(f_0)$  is constant in  $z$  and hence solves the equation

$$f_t + z^k e_k f = 0. \quad (3.60)$$

It is now easy to see that our assumptions imply that such  $f$  is constant also in  $a$ .

### 3.4 A calculation for a non-compact group

We now consider the situation in the previous subsection for the special case  $G = \mathbf{R}^n$  and a one-dimensional manifold  $Z \subset \mathfrak{g} \sim \mathbf{R}^n$ . In other words,  $Z$  will be an analytic curve in  $\mathbf{R}^n$ . We will see that analyticity is not really needed for the calculation below, but we keep it as an assumption, so that we have the Hörmander condition for the Fokker-Planck equation under the assumptions of Theorem 3.2. We will assume that  $Z$  is equipped with a measure  $m$ , the density of which is also analytic with respect to the parameter which gives an analytic parametrization of  $Z$ . We will re-parametrize  $Z$  so that it is given by an analytic periodic function

$$\gamma: R \rightarrow Z \subset \mathbf{R}^n \quad (3.61)$$

with minimal period  $l$  and, in addition, the measure (as measured by  $m$ ) of a segment on the curve between  $\gamma(s_1)$  and  $\gamma(s_2)$  for some  $0 \leq s_1 < s_2 < l$  will be given by  $s_2 - s_1$ . Sometimes we will also write

$$\gamma(s) = z(s), \quad (3.62)$$

with slight abuse of notation which will hopefully not cause any confusion. In this special case the Fokker-Planck equation discussed in the previous section, written in the variables  $a = (a^1, \dots, a^n) \in G$  and  $s$  (which parametrizes  $Z$ ), is

$$f_t + z^k(s) \frac{\partial f}{\partial a^k} = \frac{\varepsilon^2}{2} \frac{\partial^2 f}{\partial s^2}, \quad (3.63)$$

where  $f = f(a^1, \dots, a^n, s, t)$  is periodic in  $s$ , with period  $l$ . The p-hull condition from Definition 3.2 is that  $Z - Z$  generates  $\mathbf{R}^n$ .



We are interested in the long-time behavior of the solutions of (3.63). We will assume that the p-hull condition is satisfied. It is easy to see that the case when the condition is not satisfied can be reduced to this case by a suitable choice of variable.<sup>6</sup>

We note that the change of variables  $a^k \rightarrow a^k - z_0^k t$  for some  $z_0 \in \mathbf{R}^n$  is equivalent to shifting  $Z$  to  $Z - z_0$ . We can therefore assume without loss of generality that

$$\int_0^l \gamma(s) ds = \int_Z z m(dz) = 0. \quad (3.64)$$

This condition enables us to write

$$\gamma(s) = \varphi''(s) \quad (3.65)$$

for some periodic (analytic)  $\varphi: \mathbf{R} \rightarrow \mathbf{R}^n$ . An important role will be played by the matrix

$$\Sigma_{kl} = \frac{1}{l} \int_0^l \varphi'_k(s) \varphi'_l(s) ds. \quad (3.66)$$

**Proposition 3.1** *Assume (3.64) (which can be always achieved by a change of variables  $a \rightarrow a - z_0 t$ ) and let  $\Sigma_{kl}$  be defined by (3.66). For any compactly supported initial density  $f_0 = f_0(a, s)$  (normalized to total mass one) the quantity*

$$a \rightarrow t^{\frac{n}{2}} \int_0^l f(\sqrt{t} a, s, t) ds \quad (3.67)$$

*converges as  $t \rightarrow \infty$  (in distribution) to the density of the normal distribution with average 0 and covariance matrix  $\frac{4}{\varepsilon^2} \Sigma_{kl}$ . In other words, the distribution of the positions of trajectories starting at time  $t$  in some compact region will approach (after re-scaling) the same distribution as the diffusion with covariance matrix  $\frac{4}{\varepsilon^2} \Sigma_{kl}$ .*

*Proof:*

We will work with the corresponding stochastic ODE

$$\begin{aligned} \dot{a} &= \gamma(s) \\ \dot{s} &= \varepsilon \dot{w}, \end{aligned} \quad (3.68)$$

where  $w(t)$  is the standard one-dimensional Wiener process starting at the origin. Our task reduces to evaluating

$$a(t) - a(0) = \int_0^t \gamma(\varepsilon w(t')) dt' = \int_0^t \varphi''(\varepsilon w(t')) dt'. \quad (3.69)$$

We will evaluate the integral by a standard procedure based on the martingale version of the central limit theorem. We only sketch the main steps. By Itô formula we have

$$\varphi(a(t)) - \varphi(a(0)) = \int_0^t \varepsilon \varphi'(\varepsilon w(t')) dw(t') + \int_0^t \frac{\varepsilon^2}{2} \varphi''(\varepsilon w(t')) dt'. \quad (3.70)$$

We re-write this as

$$\frac{1}{\sqrt{t}} \int_0^t \gamma(\varepsilon w(t')) dt' = \int \frac{2}{\varepsilon \sqrt{t}} \varphi'(w(t')) (-dw(t')) - \frac{2}{\varepsilon^2 \sqrt{t}} (\varphi(a(t)) - \varphi(a(0))) \quad (3.71)$$

---

<sup>6</sup>Here and below this is of course meant only in the context of the example we are considering in this subsection.

The last term on the right clearly approaches zero for  $t \rightarrow \infty$ , as  $\varphi$  is bounded. The key point now is to use a martingale version of the central limit theorem (such as, for example Theorem 3.2, page 58 in [9]) to get a good asymptotics for the integral on the right. The covariance matrix for that integral generated along a trajectory  $w(t')$  is

$$\frac{4}{\varepsilon^2 t} \int_0^t \varphi'_k(\varepsilon w(t')) \varphi'_l(\varepsilon w(t')) dt'. \quad (3.72)$$

For large times  $t'$  the distribution of the variable  $\varepsilon w(t')$  taken mod  $l$  will be approaching the uniform distribution in  $[0, l)$  and therefore it is not hard to see that for the purposes of our calculation we can replace the random quantity (3.72) by a deterministic quantity given by

$$\frac{4}{\varepsilon^2 l} \int_0^l \varphi'_k(s) \varphi'_l(s) ds = \frac{4}{\varepsilon^2} \Sigma_{kl}. \quad (3.73)$$

The claim of the proposition now essentially follows from the central limit theorem.

### Acknowledgement

We thank Jonathan Mattingly for an illuminating discussion.

The research was supported in part by grants DMS 1362467 and DMS 1159376 from the National Science Foundation.

## References

- [1] Arnold, V. I., *Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits*, Ann. Inst. Fourier (Grenoble) 16 1966 fasc. 1, 319–361.
- [2] Arnold, V. I., Khesin, B. A., *Topological Methods in Hydrodynamics*, Springer, 1998.
- [3] Bismut, J.-M., *Mecanique Aléatoire*, Lecture Notes in Math., vol. 866, Springer 1981.
- [4] Einstein, A., *Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen*, Annalen der Physik, 322 (8), 549–560.
- [5] Jurdjevic, V., *Geometric Control Theory*, Cambridge Studies in Advanced Mathematics vol. 52, 1997
- [6] Hairer, M., *On Malliavin's proof of Hörmander's theorem*, Bull. Sci. Math. 135 (2011), no. 6-7, 650–666
- [7] Hairer, M., *Ergodic theory for stochastic PDEs*, Lecture notes from the LAM-EPSC short course held in July 2008.
- [8] Hairer, M., Mattingly, J. C., *Ergodicity of the 2D Navier-Stokes equations with degenerate stochastic forcing*, Ann. of Math. (2) 164 (2006), no. 3, 993–1032.
- [9] Hall, P., Heyde, C.C., *Martingale Limit Theory and its Application*, Academic Press, 1980.

- [10] Herzog, D. P., Mattingly, J. C., *A practical criterion for positivity of transition densities*, Nonlinearity, 28 (2015) 2823–2845.
- [11] Hochgerner, S., Ratiu, T., *Geometry of non-holonomic diffusion*, arXiv:1204.6438.
- [12] Hörmander, L., *Hypoelliptic second order differential equations*, Acta Math. 119, 1967, 147–171.
- [13] Khasminskii, R., *Stochastic stability of differential equations*, second edition. Stochastic Modelling and Applied Probability, 66. Springer, 2012.
- [14] Kuksin, S. B., *Randomly forced nonlinear PDEs and statistical hydrodynamics in 2 space dimensions. Zurich Lectures in Advanced Mathematics*, European Mathematical Society (EMS), Zürich, 2006.
- [15] Marsden, J., Weinstein, A., *Coadjoint orbits, vortices, and Clebsch variables for incompressible fluids*, Order in chaos (Los Alamos, N.M., 1982), Phys. D 7 (1983), no. 1-3, 305–323.
- [16] Stroock, D. W., Varadhan, S. R. S., *On the support of diffusion processes with applications to the strong maximum principle*, Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. III: Probability theory, pp. 333–359.
- [17] Villani, C., *Hypocoercivity*, Mem. Amer. Math. Soc. 202 (2009), no. 950.